

What is claimed is:

1. A database comprising information on amino acid sequences of proteins with one or more known biological functions, and further comprising information on importance scores regarding appearance of said biological functions added for each amino acid residue constituting said amino acid sequences.

2. The database according to claim 1 which is utilized to predict function of a protein with unknown biological function based on homology of amino acid sequences.

3. The database according to claim 1 ~~or claim 2~~ which is prepared by using amino acid sequences of proteins for which information on protein steric structures is available as the information on amino acid sequences of proteins with known biological functions.

4. The database according to ^{claim 1} ~~any one of claims 1 through 3~~ which is stored in a storage medium.

5. A method of preparing an alignment of a protein stored in the database according to ^{claim 1} ~~any one of claims 1 through 3~~ and a polypeptide with unknown biological function which comprises the steps of:

calculating a homology score to the coincidence of each constituent amino acid under consideration of the importance score for the appearance of a biological function, and

preparing an alignment representing homology of sites where said importance is high.

6. The method according to claim 5, which comprises the step of searching correspondence with high homology concerning the protein in the database and the target protein by using group sequences comprising two or more continuous amino acid residues of high importance with regard to the appearance of the biological function.

7. The method according to claim 5 ~~or claim 6~~, which comprises the step of obtaining a final score of homology from the alignment regarding one of the proteins in the database and the target protein.

8. The method according to claim 7, which comprises the step of predicting one or more proteins most homologous to the target protein with regard to the biological function based on the final scores for all the proteins in the database.